

Real zeroes of random polynomials, I Flip-invariance, Turán's lemma, and the Newton-Hadamard polygon

Ken Söze *

To Ildar Ibragimov with admiration

Abstract

We show that with high probability the number of real zeroes of a random polynomial is bounded by the number of vertices on its Newton-Hadamard polygon times the cube of the logarithm of the polynomial degree. A similar estimate holds for zeroes lying on any curve in the complex plane, which is the graph of a Lipschitz function in polar coordinates. The proof is based on the classical Turán lemma.

1 Introduction

This work is motivated by the following question attributed to Larry Shepp: Let

$$P(z) = \sum_{k=0}^n \lambda_k z^k$$

be a random polynomial of degree $n \geq 2$ with independent identically distributed random coefficients λ_k . *Is it true that the expected number of real zeroes of P is bounded by $C \log n$?* Since the classical work of Mark Kac [6], for many “decent” distributions of the coefficients, it has been proven by Erdős and Offord [2], Logan and Shepp [9], Ibragimov and Maslova [4, 5], Shepp and Farahmand [14] (by no means is this list complete). Here, we are interested in a bound valid for all distributions. Several years ago, Ibragimov and Zaporozhets [3] proved that for any distribution of the coefficients, the expected number of real zeroes is $o(n)$ as $n \rightarrow \infty$. Later, in works that remained unpublished, this was independently improved by Kabluchko and Zaporozhets and by Krishnapur and Zeitouni to $O(\sqrt{n})$. In the opposite direction, Zaporozhets [15] constructed an example of a distribution wherein the mean number of real zeroes remains bounded as $n \rightarrow \infty$.

In this work we suggest two approaches to this question. The first one, presented in this part, is based on tools from harmonic and complex analysis (Turán's lemma and Jensen's formula). In the case when the coefficients of P are independent and identically distributed, it gives a bound $C \log^4 n$, which is weaker than the estimate we prove in Part II. On the other hand, the approach of Part I needs less restrictive condition (which we call “the property (Θ) ”) than independence and identical distribution of the coefficients. Assuming the property (Θ) , we show that, with high probability, the number of real zeroes of P is bounded by $C V(P) \log^3 n$ where $V(P)$ is the number of vertices on the Newton-Hadamard polygon of P . It also gives the same upper bound for the number of zeroes of P on any curve in the complex plane, which is the graph of a Lipschitz function in polar coordinates.

The second approach, which we shall present in Part II, is based on Descartes' rule of sign changes and on a new anti-concentration estimate for random permutations of large order, which

*290W 232nd Str, Apt 4b, Bronx, NY 10463, USA; sozeken65@gmail.com

might be of independent interest. Both approaches may be viewed as further development of the techniques introduced in the pioneering work of Littlewood and Offord [8].

2 Main results

2.1 Key definitions

We start with three definitions needed to formulate our results. In what follows, P always stands for a polynomial of degree n with, generally speaking, complex-valued coefficients λ_k .

2.1.1 The number of vertices on the Newton-Hadamard polygon

We denote by $V(P)$ the number of vertices on the graph of the convex polygonal function

$$t \mapsto \max_{0 \leq k \leq n} (\log |\lambda_k| + kt), \quad t \in \mathbb{R}.$$

Although we will not use it, it is not difficult to see that equivalently $V(P)$ can be defined as the number of vertices on the Newton-Hadamard polygon, which is the upper envelope of convex functions φ such that $\varphi(k) \leq -\log |\lambda_k|$, $0 \leq k \leq n$ (in other words, the lower boundary of the convex hull of $n+1$ vertical rays $\{(k, y) : -\log |\lambda_k| \leq y < +\infty, 0 \leq k \leq n\}$). For more on this, see [12, Chapter IV, Problem 41].

2.1.2 The Lipschitz curves

By Γ we denote an arbitrary curve defined in polar coordinates by

$$\Gamma = \{z = re^{i\theta} : \theta = \theta(r), 0 \leq r < \infty\}.$$

If

$$|\theta(r_1) - \theta(r_2)| \leq L \left| \log \frac{r_1}{r_2} \right|,$$

then we call Γ an L -Lipschitz curve. We denote by $N(\Gamma; P)$ the number of zeroes of P on Γ (counted with multiplicities).

2.1.3 Flips of the coefficients

Let λ' and λ'' be \mathbb{C}^{n+1} -valued random variables defined on the same probability space and having the same distribution. For $k \in \{0, 1, \dots, n\}$, we put $\lambda_k^+ = \lambda'_k$ and $\lambda_k^- = \lambda''_k$, and then, for any $(n+1)$ -tuple of signs $\sigma \in \{+, -\}^{n+1}$, let $\lambda^\sigma = (\lambda_0^{\sigma_0}, \lambda_1^{\sigma_1}, \dots, \lambda_n^{\sigma_n})$. We say that the joint law of λ' and λ'' is *flip-invariant* if the random variables $\{\lambda^\sigma\}_{\sigma \in \{+, -\}^{n+1}}$ are equidistributed.

2.1.4 The property (Θ)

Here, we introduce our assumption on the distribution of the coefficients $\lambda \in \mathbb{C}^{n+1}$ of the polynomial P . We say that the coefficients of the random polynomial P possess property (Θ) if there exist random variables λ' and λ'' equidistributed with λ whose joint law is flip-invariant and such that, for some $a \in \mathbb{C}$ and for each $k \in \{0, 1, \dots, n\}$,

$$|\lambda_k^{\sigma_k} - \lambda_k^{-\sigma_k}| \geq \frac{1}{2} [|\lambda_k^{\sigma_k} - a| + |\lambda_k^{-\sigma_k} - a|] \quad \text{a.s. .}$$

Note that for our purposes, it would suffice to have this inequality with any constant $\kappa > 0$ instead of $\frac{1}{2}$. In the examples, which we will bring below, this condition holds with the value $\kappa = \frac{1}{2}$. To simplify our notation, we decided to fix this value of κ .

2.1.5 Three examples of distributions with property (Θ)

Symmetric distributions. For $k \in \{0, 1, \dots, n\}$, denote by $\tau_k: \mathbb{C}^{n+1} \rightarrow \mathbb{C}^{n+1}$ the map, which maps $w_k \mapsto -w_k$ and keeps fixed the rest of coordinates of $w \in \mathbb{C}^{n+1}$. Suppose that, for each $k \in \{0, 1, \dots, n\}$, $\tau_k \circ \lambda$ has the same distribution as λ . Then the distribution of λ has property (Θ) with $a = 0$, $\lambda' = \lambda$, and $\lambda'' = -\lambda$.

Note that in this example we do not assume independence of λ_k 's.

Complex-valued independent identically distributed random variables $\lambda_0, \lambda_1, \dots, \lambda_n$. Denote by ζ the common distribution of λ_k 's. We need to produce two random variables ζ^\pm having the same distribution as ζ and such that, for some $a \in \mathbb{C}$,

$$|\zeta^+ - \zeta^-| \geq \frac{1}{2} [|\zeta^+ - a| + |\zeta^- - a|].$$

We first assume that the probability space Ω is a union of $2N$ atoms ω_i having the same probability $\frac{1}{2N}$. Then the general case will follow by approximation¹.

Let

$$d = \max_{1 \leq i < j \leq 2N} |\zeta(\omega_i) - \zeta(\omega_j)|$$

be the diameter of the point configuration $\{\zeta(\omega_1), \dots, \zeta(\omega_{2N})\}$ in \mathbb{C} . Pick up from this configuration a pair of points with the maximal distance. Without loss of generality, assume that they correspond to the atoms ω_{2N} and ω_{2N-1} , that is, $d = |\zeta(\omega_{2N-1}) - \zeta(\omega_{2N})|$. Then consider the remaining point configuration and repeat the procedure. At the last N -th step we are left with two points $\zeta(\omega_1)$ and $\zeta(\omega_2)$. Then denote by a the center of the line segment that connects these two points, that is, $a = \frac{1}{2} (\zeta(\omega_1) + \zeta(\omega_2))$.

By construction, for each $1 \leq i \leq N$, the point a lies at distance at most $|\zeta(\omega_{2i-1}) - \zeta(\omega_{2i})|$ from each of the two points $\zeta(\omega_{2i-1})$, $\zeta(\omega_{2i})$. Hence,

$$|\zeta(\omega_{2i-1}) - \zeta(\omega_{2i})| \geq \frac{1}{2} [|\zeta(\omega_{2i-1}) - a| + |\zeta(\omega_{2i}) - a|].$$

It remains to let $\zeta^+ = \zeta$, and

$$\zeta^-(\omega_{2i-1}) = \zeta(\omega_{2i}), \quad \zeta^-(\omega_{2i}) = \zeta(\omega_{2i-1}), \quad 1 \leq i \leq N.$$

¹ Indeed, take a sequence of random variables (ζ_N) that converges in distribution to ζ and such that ζ_N attains $2N$ values (not necessarily distinct ones) with probability $\frac{1}{2N}$ each. Let (ζ_N^+, ζ_N^-) be a pair of random variables defined on the same probability space as ζ_N , equidistributed with ζ_N and such that, for some $a_N \in \mathbb{C}$,

$$|\zeta_N^+ - \zeta_N^-| \geq \frac{1}{2} [|\zeta_N^+ - a_N| + |\zeta_N^- - a_N|]. \quad (*)$$

Since ζ_N converge to ζ in distribution, the sequence of laws of ζ_N is tight. Then the sequence of joint laws of pairs (ζ_N^+, ζ_N^-) is tight as well, and we can choose a subsequence $(\zeta_{N_j}^+, \zeta_{N_j}^-)$ that converges in distribution to a pair of random variables (ζ^+, ζ^-) defined on the same probability space as ζ and equidistributed with ζ .

Furthermore, by tightness of the sequence of laws of (ζ_N^+, ζ_N^-) , we can choose a large positive constant L so that, for every N , $\mathbb{P}\{|\zeta_N^+ - \zeta_N^-| > L\} < \frac{1}{2}$. Therefore, on an event of probability at least $\frac{1}{2}$,

$$|a_N| \leq |\zeta_N^+ - a_N| + |\zeta_N^- - a_N| + |\zeta_N^+ - \zeta_N^-| \leq 3|\zeta_N^+ - \zeta_N^-| \leq 3L.$$

Since both a_N and L are non-random, it follows that $|a_N| \leq 3L$. Then, extracting from (a_{N_j}) a convergent subsequence, denoting by a its limit, and applying $(*)$ with $N = N_j$, $j \rightarrow \infty$, we get the result.

Real-valued independent random variables $\lambda_0, \lambda_1, \dots, \lambda_n$ which have a common median.

Arguing similarly to the previous example, we construct the coefficients λ_k^\pm equidistributed with λ_k and satisfying $\lambda_k^+ + \lambda_k^- = 2a$, $0 \leq k \leq n$, where a is the common median for $\lambda_0, \lambda_1, \dots, \lambda_n$.

Note that in this example we have not assumed that the coefficients λ_k are identically distributed.

2.1.6 A technical assumption

To avoid degeneracies, in what follows, we always assume that the coefficients of the random polynomial P satisfy

$$\mathbb{P}\{\lambda_0 = 0\} = \mathbb{P}\{\lambda_n = 0\} = 0. \quad (1)$$

That is, P does not vanish at the origin and the degree of P does not drop. This condition can be dropped at the cost of a somewhat longer wording of the main result.

2.2 The main theorem

At last, we are ready to state the main result of this note:

Theorem 1. *Let P be a random polynomial of degree $n \geq 2$ with coefficients having the property (Θ) and satisfying the non-degeneracy condition (1). Let $L \geq 0$ and $A > 0$. Then, with probability at least $1 - n^{-A}$, we have*

$$\sup\{N(\Gamma; P) : \Gamma \text{ is } L\text{-Lipschitz}\} \leq C(A, L) V(P) \log^3 n.$$

Here, $C(A, L)$ is a positive value that depends only on the parameters A and L .

Note that there is no hope for a similarly strong non-random estimate: a construction, which goes back to Bloch and Pólya [1], allows one to construct a polynomial P of any degree $n \geq 2$ with $V(P) = 2$ and with at least $\sqrt{n/\log n}$ positive zeroes.

2.3 A corollary for the case of i.i.d. coefficients

As an almost immediate corollary, we obtain

Corollary 2. *Suppose that the coefficients of P are independent identically distributed random variables satisfying the non-degeneracy condition (1). Then*

$$\sup\{\mathbb{E}[N(\Gamma; P)] : \Gamma \text{ is } L\text{-Lipschitz}\} \leq C(L) \log^4 n, \quad n \geq 2.$$

In particular, $\mathbb{E}[N(\mathbb{R}; P)] \leq C \log^4 n$ with a positive numerical constant C . As we have already mentioned, the latter estimate will be improved in Part II by a different technique.

Proof. We use Theorem 1 with $A = 1$. Since the total number of zeroes of P on Γ cannot exceed n , a set of probability n^{-1} can contribute to the expectation $\mathbb{E}[N(\Gamma; P)]$ by at most 1. Therefore,

$$\sup\{\mathbb{E}[N(\Gamma; P)] : \Gamma \text{ is } L\text{-Lipschitz}\} \leq C(L) \log^3 n \cdot \mathbb{E}[V(P)].$$

To estimate the mean $\mathbb{E}[V(P)]$, we note that $V(P)$ equals the cardinality of the set of indices $\nu \in \{0, 1, \dots, n\}$ such that, for some $r \in (0, \infty)$,

$$\nu \text{ is the largest index satisfying } |\lambda_\nu| r^\nu = \max_{0 \leq k \leq n} |\lambda_k| r^k, \quad (2)$$

If (2) holds for some $r \in (0, 1)$, then

$$|\lambda_\nu| > |\lambda_k| \quad \text{for each } k \in \{0, 1, \dots, \nu - 1\}.$$

By symmetry, the probability of this event does not exceed $\frac{1}{\nu+1}$. Similarly, if (2) holds for some $r \in [1, \infty)$, then

$$|\lambda_\nu| > |\lambda_k| \quad \text{for each } k \in \{\nu+1, \nu+2, \dots, n\},$$

and the probability of this event is $\leq \frac{1}{n-\nu+1}$. Thus,

$$\mathbb{E}[V(P)] \leq 2\left(1 + \frac{1}{2} + \dots + \frac{1}{n}\right) \leq C \log n, \quad n \geq 2,$$

proving the corollary. \square

2.4 A probabilistic lower bound for a random polynomial on an arc

The following result is the main tool needed for the proof of Theorem 1. Likely, it may be of independent interest. Put

$$S(r, P) = \sum_{k=0}^n |\lambda_k| r^k.$$

Theorem 3. *Let P be a random polynomial of degree $n \geq 2$ with coefficients having the property (Θ) . Let $m \in \mathbb{N}$, let $r > 0$, and let $I \subset \mathbb{R}$ be an interval of length at most 2π . Then, for some positive numerical constant c ,*

$$\mathbb{P}\left\{\max_{\theta \in I} |P(re^{i\theta})| \leq n^{-2}(c|I|)^{6m} S(r, P)\right\} \leq 2^{-m}.$$

The proof of this theorem will be given in Section 3.

2.5 The reduction principle

Our starting point is the following claim:

Lemma 4. *Suppose that the coefficients of the random polynomial P possess the property (Θ) . Then, for any Borel set $\Lambda \subset \mathbb{C}^{n+1}$, we have*

$$\mathbb{P}\{\lambda \in \Lambda\} \leq \sup \mathbb{P}\{v \in \Lambda\},$$

where the supremum is taken over all random variables $v: \{+, -\}^{n+1} \rightarrow \mathbb{C}^{n+1}$ of the form $v^\sigma = (v_0^{\sigma_0}, v_1^{\sigma_1}, \dots, v_n^{\sigma_n})$ such that the random variables v_k are independent, take the values v_k^\pm with probability $\frac{1}{2}$ and, for some $a \in \mathbb{C}$,

$$|v_k^+ - v_k^-| \geq \frac{1}{2} [|v_k^+ - a| + |v_k^- - a|], \quad k \in \{0, 1, \dots, n\}. \quad (3)$$

It is worth noting that for independent real-valued random variables, this reduction was used already by Kolmogorov in [7], where he proved a slightly weaker version of what is called nowadays the Kolmogorov-Rogozin concentration inequality.

Proof. Let Ω be the underlying probability space of λ' and λ'' in the definition of flip-invariance, let $\tilde{\Omega} = \Omega \times \{+, -\}^{n+1}$ be the product space with the uniform distribution over all sign sequences $\sigma = (\sigma_0, \sigma_1, \dots, \sigma_n)$, and let $\lambda^\sigma \stackrel{\text{def}}{=} (\lambda_0^{\sigma_0}, \lambda_1^{\sigma_1}, \dots, \lambda_n^{\sigma_n})$. Then $\lambda^\sigma: \tilde{\Omega} \rightarrow \mathbb{C}^{n+1}$ and, for each $\sigma \in \{+, -\}^{n+1}$, the random variables λ^σ and λ are equidistributed. Therefore,

$$\mathbb{P}^\Omega\{\lambda \in \Lambda\} = \mathbb{P}^{\Omega \times \{+, -\}^{n+1}}\{\lambda^\sigma \in \Lambda\} \leq \operatorname{ess\,sup}_{\omega \in \Omega} \mathbb{P}^{\{+, -\}^{n+1}}\{\lambda^\sigma(\omega) \in \Lambda\}.$$

It remains to observe that, for a.e. $\omega \in \Omega$, the random variable $v^\sigma = \lambda^\sigma(\omega)$ satisfies (3) with the same value a as in the condition (Θ) . Hence, the essential supremum on the RHS does not exceed the supremum in the conclusion of the lemma. \square

Thus, it suffices to prove Theorems 1 and 3 for a special class of random polynomials. Hence, in what follows, we assume that:

- (a) the underlying probability space is $\{+, -\}^{n+1}$ with the uniform distribution over sign sequences, and, as above, we denote the elements of this space by $\sigma = (\sigma_0, \sigma_1, \dots, \sigma_n)$;
- (b) (λ_k^\pm) are $2n + 2$ complex numbers, a is a complex number, and for each $k \in \{0, 1, \dots, n\}$,

$$|\lambda_k^+ - \lambda_k^-| \geq \frac{1}{2} [|\lambda_k^+ - a| + |\lambda_k^- - a|];$$

- (c) the random variables λ_k are independent and λ_k takes the values λ_k^\pm with probability $\frac{1}{2}$ each.

3 Proof of Theorem 3

The main ingredient of the proof of Theorem 3 is Turán's lemma [10, Chapter 5, Lemma 1] (see also [11, Chapter 1]):

Lemma 5. *Let*

$$p(t) = \sum_{k=1}^m a_k e^{i\ell_k t}, \quad a_k \in \mathbb{C}, \quad \ell_k \in \mathbb{Z}, \ell_k \neq \ell_j \text{ for } k \neq j.$$

Then for every interval $I \subset \mathbb{R}$ of length at most 2π ,

$$\max_I |p| \geq (b|I|)^{m-1} \sum_{k=1}^m |a_k|$$

with a positive numerical constant b .

Note that the conclusion of this lemma is usually stated in the form

$$\max_I |p| \geq (b|I|)^{m-1} \max_{[-\pi, \pi]} |p|.$$

Since

$$\sum_{k=1}^m |a_k| \leq \sqrt{m} \left(\sum_{k=1}^m |a_k|^2 \right)^{1/2} = \sqrt{m} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |p(t)|^2 dt \right)^{1/2} \leq \sqrt{m} \max_{[-\pi, \pi]} |p|,$$

the version we will be using readily follows from the usual one.

3.1 The case of few large coefficients

Given $m \in \mathbb{N}$ and $r > 0$, we assume that, for some $a \in \mathbb{C}$,

$$\#\{k: |\lambda_k - a|r^k \geq \delta S(r; P)\} \leq 2m \tag{4}$$

and show that for every interval $I \subset \mathbb{R}$ of length at most 2π ,

$$\max_{\theta \in I} |P(re^{i\theta})| \geq cn^{-1} (b|I|)^{4m+1} S(r; P) \tag{5}$$

provided that $\delta = c_1 n^{-2} (b|I|)^{4m+1}$ with a sufficiently small constant c_1 .

3.1.1 The polynomial \bar{P}

Put $\bar{P}(z) = (1 - z)P(z)$. We need this polynomial to get rid of the dependence on the value of a . Note that when $a = 0$ this polynomial is not needed.

Claim 6.

$$S(r; \bar{P}) \geq \frac{1+r}{2(n+1)} S(r; P), \quad 0 < r < \infty. \quad (6)$$

Proof. First, assume that $0 < r \leq 1$. Then

$$\begin{aligned} S(r; \bar{P}) &= |\lambda_0| + \sum_{k=1}^n |\lambda_k - \lambda_{k-1}| r^k + |\lambda_n| r^{n+1} \\ &\geq \frac{1}{n+1} \left[(n+1)|\lambda_0| + \sum_{k=1}^n (n+1-k) |\lambda_k - \lambda_{k-1}| r^k \right] \\ &\geq \frac{1}{n+1} \left[(n+1)|\lambda_0| + \sum_{k=1}^n (n+1-k) |\lambda_k| r^k - \sum_{k=1}^n (n+1-k) |\lambda_{k-1}| r^k \right] \\ &= \frac{1}{n+1} \left[(n+1)|\lambda_0| + \sum_{k=1}^{n-1} ((n+1-k) - (n-k)r) |\lambda_k| r^k + |\lambda_n| r^n - n|\lambda_0| r \right] \\ &= \frac{1}{n+1} \left[\sum_{k=0}^{n-1} ((n+1-k) - (n-k)r) |\lambda_k| r^k + |\lambda_n| r^n \right]. \end{aligned}$$

For $0 < r \leq 1$, we have $(n+1-k) - (n-k)r \geq 1$. Thus, the RHS of the previous estimate is

$$\geq \frac{1}{n+1} \sum_{k=0}^n |\lambda_k| r^k \stackrel{0 < r \leq 1}{\geq} \frac{1+r}{2(n+1)} S(r; P).$$

Now, let $1 \leq r < \infty$. Then

$$\begin{aligned} S(r; \bar{P}) &= |\lambda_0| + \sum_{k=1}^n |\lambda_k - \lambda_{k-1}| r^k + |\lambda_n| r^{n+1} \\ &\geq \frac{1}{n+1} \left[\sum_{k=1}^n k (|\lambda_{k-1}| r^k - |\lambda_k| r^k) + (n+1) |\lambda_n| r^{n+1} \right] \\ &= \frac{1}{n+1} \sum_{k=0}^n ((k+1)r - k) |\lambda_k| r^k. \end{aligned}$$

Since $r \geq 1$, we have $(k+1)r - k \geq r$, and therefore, the RHS of the previous estimate is

$$\geq \frac{r}{n+1} \sum_{k=0}^n |\lambda_k| r^k \stackrel{r \geq 1}{\geq} \frac{1+r}{2(n+1)} S(r; P),$$

proving the claim. □

3.1.2 Proof of the lower bound (5) assuming (4)

First, we observe that

$$\#\{k: |\bar{\lambda}_k| r^k \geq 2\delta(1+r)S(r; P)\} \leq 4m+2, \quad (7)$$

where $\bar{\lambda}_k$ are coefficients of the polynomial \bar{P} . Indeed,

$$\bar{P}(z) = \lambda_0 + \sum_{k=1}^n (\lambda_k - \lambda_{k-1})z^k - \lambda_n z^{n+1}.$$

Suppose that for some $k \in \{1, \dots, n\}$ and $\delta > 0$,

$$|\lambda_k - \lambda_{k-1}|r^k \geq 2\delta(1+r)S(r; P).$$

Then

$$|\lambda_k - a|r^k + |\lambda_{k-1} - a|r^k \geq 2\delta(1+r)S(r; P).$$

That is, at least one of the following estimates holds: either $|\lambda_k - a|r^k \geq \delta(1+r)S(r; P)$, or $|\lambda_{k-1} - a|r^k \geq \delta(1+r)S(r; P)$, proving (7).

Now, we split the polynomial \bar{P} into large and small parts. The small part \bar{P}_{sm} will consists of the terms $\bar{\lambda}_k r^k$ with

$$|\bar{\lambda}_k|r^k \leq 2\delta(1+r)S(r; P).$$

The rest goes to the large part \bar{P}_{1a} , which is a sum of at most $4m+2$ terms. Using Turán's lemma, we get

$$\begin{aligned} \max_{\theta \in I} |P(re^{i\theta})| &\geq (1+r)^{-1} \max_{\theta \in I} |\bar{P}(re^{i\theta})| \\ &\geq (1+r)^{-1} \left[\max_{\theta \in I} |\bar{P}_{1a}(re^{i\theta})| - \max_{\theta \in I} |\bar{P}_{\text{sm}}(re^{i\theta})| \right] \\ &\geq (1+r)^{-1} \left[(b|I|)^{4m+1} S(r; \bar{P}_{1a}) - S(r; \bar{P}_{\text{sm}}) \right] \\ &\geq (1+r)^{-1} \left[(b|I|)^{4m+1} (S(r; \bar{P}) - (n+1)2\delta(1+r)S(r; P)) \right. \\ &\quad \left. - (n+1)2\delta(1+r)S(r; P) \right] \\ &\stackrel{(6)}{\geq} \left[(b|I|)^{4m+1} \left(\frac{1}{2(n+1)} - 2(n+1)\delta \right) - 2(n+1)\delta \right] S(r; P) \\ &\geq \left[(b|I|)^{4m+1} \left(\frac{1}{4n} - 4n\delta \right) - 4n\delta \right] S(r; P). \end{aligned}$$

Choosing $\delta = c_1 n^{-2} (b|I|)^{4m+1}$ with a sufficiently small constant c_1 , we see that the RHS of the previous estimate is $\geq c_2 n^{-1} (b|I|)^{4m+1} S(r; P)$, proving (5). \square

3.2 The dangerous configurations are rare

Fix an interval $I \subset \mathbb{R}$ of length at most 2π and fix δ as above. Taking into account what we have just proven, we see that in order to prove Theorem 3, we need to estimate the number of sign sequences $\sigma \in \{+, -\}^{n+1}$ such that *there exist at least $2m$ ($\leq n$) indices k satisfying*

$$|\lambda_k^{\sigma_k} - a|r^k \geq \delta S(r; P) \tag{8}$$

but still

$$\max_{\theta \in I} |P(re^{i\theta})| \leq \delta_1 S(r; P) \tag{9}$$

with some positive $\delta_1 \ll \delta$ to be chosen momentarily. We call the corresponding sequence of signs σ *dangerous* and aim to show that the number of dangerous sequences does not exceed 2^{n+1-m} .

Take any dangerous sign sequence σ and an m -element subset of the set of “large coefficients” that appear in condition (8), and flip all the signs σ_k corresponding to this m -element subset. Running over all possible m -elements subsets of the set of “large coefficients” of a given dangerous sign sequence σ , we obtain at least $\binom{2m}{m} \geq 2^m$ different sign sequences. Claim 7 (given few lines below) will yield that *all new sign sequences obtained from all dangerous sign sequences σ are different, provided that the parameter δ_1 is chosen as*

$$\delta_1 = \frac{1}{4} \delta (b|I|)^{2m-1}. \quad (10)$$

Therefore, with the choice of the parameters as in (10), *the total number of all dangerous sign sequences multiplied by $\binom{2m}{m}$ cannot exceed 2^{n+1}* . At the same time, for any non-dangerous σ , we automatically have

$$\max_{\theta \in I} |P(re^{i\theta})| > \frac{c_1}{4} n^{-2} (b|I|)^{4m+1} \cdot (b|I|)^{2m-1} S(r; P) > n^{-2} (c|I|)^{6m} S(r; P).$$

Therefore, Theorem 3 follows if we prove the following claim:

Claim 7. *Let $\sigma \in \{+, -\}^{n+1}$ be any sign sequence. Suppose that there exist two different m -element subsets $U_1, U_2 \subset \{0, 1, 2, \dots, n\}$, $U_1 \neq U_2$, so that the sets of flips corresponding to U_1 and U_2 turn σ into a dangerous sign sequence with all coefficients corresponding to flipped signs becoming “large” as in condition (8). Then the parameter δ_1 cannot be as small as in (10).*

Proof. Once again, we will rely on Turán’s lemma. We fix the sign sequence σ and denote by σ^1, σ^2 the flipped sign sequences, i.e.,

$$\sigma_k^j = \begin{cases} -\sigma_k & \text{for } k \in U_j, \\ \sigma_k & \text{for } k \notin U_j. \end{cases}$$

By P_j , $j = 1, 2$, we denote the corresponding polynomials. Choosing $k_1 \in U_1 \setminus U_2$ and $k_2 \in U_2 \setminus U_1$, we have

$$\max_{\theta \in I} |P_j(re^{i\theta})| \stackrel{(9)}{\leq} \delta_1 S(r; P_j) \stackrel{(8)}{\leq} \frac{\delta_1}{\delta} |\lambda_{k_j}^{\sigma_{k_j}^j} - a| r^{k_j}, \quad j = 1, 2.$$

Therefore,

$$\max_{\theta \in I} |(P_1 - P_2)(re^{i\theta})| \leq \frac{\delta_1}{\delta} \left[|\lambda_{k_1}^{\sigma_{k_1}^1} - a| r^{k_1} + |\lambda_{k_2}^{\sigma_{k_2}^2} - a| r^{k_2} \right]. \quad (11)$$

On the other hand, the difference $P_1 - P_2$ has at most $2m$ terms:

$$(P_1 - P_2)(z) = \sum_{k \in U_1 \triangle U_2} (\lambda_k^{\sigma_k^1} - \lambda_k^{\sigma_k^2}) z^k$$

where, as usual, \triangle denotes the symmetric difference. For $k \in U_1 \triangle U_2$, we have $\sigma_k^2 = -\sigma_k^1$. Then, by assumption (b) in Section 2.5,

$$|\lambda_k^{\sigma_k^1} - \lambda_k^{\sigma_k^2}| \geq \frac{1}{2} \left[|\lambda_k^{\sigma_k^1} - a| + |\lambda_k^{\sigma_k^2} - a| \right], \quad k \in U_1 \triangle U_2.$$

In particular, this holds for $k = k_1, k_2$. Therefore, the RHS of (11) is

$$\leq \frac{2\delta_1}{\delta} \left[|\lambda_{k_1}^{\sigma_{k_1}^1} - \lambda_{k_1}^{\sigma_{k_1}^2}| r^{k_1} + |\lambda_{k_2}^{\sigma_{k_2}^1} - \lambda_{k_2}^{\sigma_{k_2}^2}| r^{k_2} \right] \leq \frac{2\delta_1}{\delta} S(r; P_1 - P_2).$$

If δ_1 is as small as in (10), this contradicts Turán’s lemma applied to $P_1 - P_2$. This proves the claim and finishes off the proof of Theorem 3. \square

4 Proof of Theorem 1

4.1 Preliminaries

4.1.1

First, we observe that it suffices to prove Theorem 1 only for zeroes of P lying in the closed unit disk $\{|z| \leq 1\}$. To get the result for the rest of the zeroes, all one needs is to consider the polynomial $P^*(z) = z^n P(z^{-1})$.

4.1.2

It will be convenient to make the exponential change of variable $z = e^{-2\pi w}$, $w = t + is$, $0 \leq t < \infty$, and to deal with the exponential polynomial

$$Q(w) = P(z) = \sum_{k=0}^n \lambda_k e^{-2\pi k w}.$$

4.1.3

Put

$$h(t) = \max_{0 \leq k \leq n} (\log |\lambda_k| - 2\pi k t), \quad H(t) = e^{h(t)}.$$

By $\nu(t)$ we denote *the central index*, that is, the largest of the indices ν , for which

$$\log |\lambda_\nu| - 2\pi \nu t \geq \log |\lambda_k| - 2\pi k t, \quad k \in \{0, 1, 2, \dots, n\}.$$

Obviously, $H(t) \leq S(e^{-2\pi t}; P) \leq (n+1)H(t)$. This will allow us to replace S by H in our estimates. The advantage of H over S is that the former has sharper transitions at the points where the central index changes its value.

4.1.4

In the new notation, Theorem 3 says that *given $t \geq 0$, given an interval $I = [s', s'']$ of length less than 1, and given a positive integer parameter m , there exists an event*

$$\Sigma(t, I, m) \subset \{+, -\}^{n+1}, \quad \text{with} \quad \mathbb{P}(\Sigma(t, I, m)) \leq 2^{-m}$$

such that for every $\sigma \in \{+, -\}^{n+1} \setminus \Sigma(t, I, m)$,

$$\max_{s \in I} |Q(t + is)| \geq n^{-2} (c|I|)^{6m} H(t). \quad (12)$$

This estimate is complemented by the obvious upper bound

$$\max_s |Q(t + is)| \leq (n+1)H(t). \quad (13)$$

4.2 The test sets and exceptional sign sequences

Our exceptional event $\Sigma \subset \{+, -\}^{n+1}$ will be a union of the events $\Sigma(t, I, m)$ taken over a certain finite sets of “test points” t and “test intervals” I . So we start by defining these sets.

4.2.1

Recall that each λ_k attains two values, consider the $2n + 2$ lines $t \mapsto \log |\lambda_k^\pm| - 2\pi kt$, $0 \leq k \leq n$. There are at most $\binom{2n+2}{2} = (n+1)(2n+1)$ points on $[0, \infty)$ where two of these functions are equal. We denote this set of points by \mathfrak{T}_0 . Then we put

$$\mathfrak{T} = \left\{ t = \frac{j}{n} : \text{dist}(t, \mathfrak{T}_0) \leq 1, j \in \mathbb{Z}_+ \right\}.$$

This will be our set of *test points* t .

Put

$$\mathfrak{S} = \left\{ s = \frac{k}{n} : 0 \leq k < n, k \in \mathbb{Z}_+ \right\}.$$

The set \mathfrak{J} of *test intervals* I will consist of all intervals centered at all the points $s \in \mathfrak{S}$, of length j/n with $1 \leq j \leq n$, $j \in \mathbb{N}$.

Claim 8. *The cardinality of the set \mathfrak{T} is $\leq Cn^3$. The cardinality of the set \mathfrak{J} is $\leq Cn^2$.*

Proof. Obvious. □

4.2.2

Now, we define the exceptional event $\Sigma \subset \{+, -\}^{n+1}$. Put

$$\Sigma(m) = \bigcup_{t \in \mathfrak{T}, I \in \mathfrak{J}} \Sigma(t, I, m)$$

where the events $\Sigma(t, I, m)$ are the same as in 4.1.4. Then, by the last claim, $\mathbb{P}(\Sigma(m)) \leq Cn^5 2^{-m}$. Now, we fix

$$m = C(A) \log n$$

with a sufficiently large value $C(A)$, and let $\Sigma = \Sigma(m)$. Then $\mathbb{P}(\Sigma) < n^{-A}$.

In the rest of the proof, we fix the sign sequence $\sigma \in \{+, -\}^{n+1} \setminus \Sigma$. We put $V = V(P)$, where $V(P)$ is the number of vertices on the Newton-Hadamard polygon introduced in Section 2.1.1.

4.3 The Whitney-type partition

4.3.1

For $t \geq 0$, the graph of the function $h(t)$ is a piece-wise linear function with at most $V + 1$ intervals of linearity. Take one of these intervals and call it \mathbb{J} . The proof of Theorem 1 needs a special partition of the interval \mathbb{J} . To construct this partition, we take L as in Theorem 1, let $L' = [L] + 4$ (as usual, $[L]$ stands for the integer part of L) and take a sequence of closed intervals with disjoint interiors starting in both directions from the test points in $\mathfrak{T} \setminus \text{int}(\mathbb{J})$ closest to the end points of \mathbb{J} so that

- the end-points of each interval of this sequence belong to the set \mathfrak{T} of test-points;
- the first $4L'$ intervals starting with each end-point of \mathbb{J} have length $\frac{2}{n}$, the next $4L'$ intervals have length $\frac{4}{n}$, the next $4L'$ have the length $\frac{8}{n}$, and so on, until we either reach length 1 or cover the middle point of \mathbb{J}

(see Fig. 1). We denote the intervals of this sequence by J and note that we used at most $CL' \log n$ intervals J per each interval \mathbb{J} .

Next, we list several properties of this construction, which will be used in the proof of Theorem 1.

4.3.2

The centers c_J of the intervals J belong to the set \mathfrak{T} of tested points. The intervals $I = [s - \frac{1}{2}|J|, s + \frac{1}{2}|J|]$, $s \in \mathfrak{S}$, belong to the set \mathfrak{J} of tested intervals.

4.3.3

By J' we denote the interval centered at c_J which is L' times longer than J . Then, by construction, if J is an interval from our partition with $|J| \geq \frac{4}{n}$, then $J' \subset \mathbb{J}$.

4.4 There are no zeroes in the strips with the untested ground

By \mathbb{J}_0 we denote the part of \mathbb{J} that remains uncovered by intervals J and call it *the untested part of \mathbb{J}* . For some intervals \mathbb{J} , the untested part \mathbb{J}_0 can be void. Let $\Pi_{\mathbb{J}_0} = \{t + is : t \in \mathbb{J}_0\}$ be the corresponding vertical strip.

Claim 9. *The exponential polynomial Q does not vanish on all vertical strips $\Pi_{\mathbb{J}_0}$.*

Proof of Claim 9: Suppose that the point t belongs to one of the intervals \mathbb{J}_0 , that is, the central index ν stays fixed on $[t - 1, t + 1]$. Thus, we actually have not only

$$\log |\lambda_\nu| - 2\pi\nu t \geq \log |\lambda_k| - 2\pi kt,$$

but also

$$\log |\lambda_\nu| - 2\pi\nu t \geq \log |\lambda_k| - 2\pi kt + 2\pi|k - \nu|.$$

Then

$$\begin{aligned} \left| \sum_{k \neq \nu} \lambda_k e^{-2\pi kt} \right| &\leq \sum_{k \neq \nu} |\lambda_k| e^{-2\pi kt} \\ &\leq |\lambda_\nu| e^{-2\pi\nu t} \sum_{k \neq \nu} e^{-2\pi|k - \nu|} = |\lambda_\nu| e^{-2\pi\nu t} \cdot 2 \sum_{k \geq 1} e^{-2\pi k} < |\lambda_\nu| e^{-2\pi\nu t}. \end{aligned}$$

Hence, Q cannot vanish on the vertical line $t + i\mathbb{R}$, and therefore, on the whole vertical strip $\Pi_{\mathbb{J}_0}$. \square

4.5 Jensen's bound for the number of zeroes of Q in the disks $\bar{D}_{J,s}$

Given interval J from our partition and $s \in \mathfrak{S}$, consider the disks

$$D_{J,s} = \{w : |w - (c_J + is)| < \frac{1}{2}L'|J|\},$$

and denote by $N(\bar{D}_{J,s}; Q)$ the number of zeroes of Q in the closed disk $\bar{D}_{J,s}$ counted with multiplicities.

Claim 10. *We have*

$$N(\bar{D}_{J,s}; Q) \leq C(A, L) \log^2 n.$$

The proof of this claim relies upon “the classical Jensen’s bound”²: *Let F be an analytic function in a disk D . Let $\frac{1}{2}D$ be the disk concentric with D but of twice smaller radius. Then the number of zeroes of F in the closed disk $\frac{1}{2}\bar{D}$ (counted with multiplicities) is*

$$\leq C \log \frac{\sup_D |F|}{\max_{\frac{1}{2}\bar{D}} |F|}.$$

Proof of Claim 10: By 4.3.3, the intervals $J \subset \mathbb{J}$ fall into two categories: either $J' \subset \mathbb{J}$, or the length of J is $\frac{2}{n}$. First, consider the intervals J from the first group, that is, assume that the central index ν stays fixed on J' . Take the function $F(w) = Q(w)e^{2\pi\nu w}$ which has the same zeroes as Q . By 4.3.2, each point c_J and each interval $I_{J,s} = [s - \frac{1}{2}|J|, s + \frac{1}{2}|J|]$ are tested. Note that $c_J + iI_{J,s} \subset D_{J,s}$. Therefore, we have the lower bound

$$\max_{\bar{D}_{J,s}} |F| \stackrel{(12)}{\geq} \max_{v \in I_{J,s}} |F(c_J + iv)| \geq n^{-2} (c|I_{J,s}|)^{6m} H(c_J) \cdot e^{2\pi\nu c_J} = n^{-2} (c|J|)^{6m} |\lambda_\nu|.$$

The matching upper bound

$$\max_{t+iv \in 2\bar{D}_{J,s}} |F(t+iv)| \stackrel{(13)}{\leq} (n+1) \max_{t \in J'} [H(t)e^{2\pi\nu t}] < 2n|\lambda_\nu|$$

is evident. Using Jensen’s bound, recalling that $|J| \geq \frac{4}{n}$ and that $m = C(A) \log n$, we get

$$N(\bar{D}_J; F) \leq Cm \log n \leq C(A) \log^2 n.$$

Now, we turn to the second case, when $|J| = \frac{2}{n}$. These intervals are so short that the function h can change only by a constant (depending on L') on J' . Indeed, let $J' = [a, b]$. Take the points $a = t_0 < t_1 < \dots < t_s = b$, so that $\nu(t_i + 0) = \nu(t_{i+1} - 0) = \nu_i$. Then

$$h(a) - h(b) = \sum_{i=0}^{s-1} [h(t_i) - h(t_{i+1})] = \sum_{i=0}^{s-1} 2\pi\nu_i [t_{i+1} - t_i] \leq 2\pi n(b - a) = 4\pi L';$$

the estimate in the opposite direction is obvious since the function h does not increase. Therefore, $H(a)/H(b) \leq e^{4\pi(L+4)}$.

Then, similarly to the first case, we take the corresponding test intervals $I_{J,s}$, note that

$$\max_{\bar{D}_{J,s}} |Q| \geq \max_{I_{J,s}} |Q| \stackrel{(12)}{\geq} n^{-2} (c|I_J|)^{6m} H(c_J),$$

and that

$$\max_{2\bar{D}_{J,s}} |Q| \stackrel{(13)}{\leq} (n+1) \max_{J'} H < 2n e^{4\pi(L+4)} H(c_J).$$

Then, applying Jensen’s bound to the function Q , and recalling that $|I_{J,s}| = |J| = \frac{2}{n}$ and that $m = C(A) \log n$, we get $N(\bar{D}_{J,s}, Q) \leq C(A, L) \log^2 n$. This proves Claim 10. \square

² For the reader’s convenience, we recall its short proof, assuming, without loss of generality, that D is the unit disk. Let a_1, \dots, a_N be zeroes of F in $\frac{1}{2}\bar{D}$ counted with multiplicities, and let

$$B_a(z) = \frac{z - a}{1 - \bar{a}z}.$$

Then $F = B_{a_1} \dots B_{a_N} G$, where the function G is analytic in D and $\sup_D |F| = \sup_D |G|$. Note that the absolute value of each factor B_{a_i} is bounded by $\frac{4}{5}$ in $\frac{1}{2}\bar{D}$. Therefore,

$$\max_{\frac{1}{2}\bar{D}} |F| \leq \left(\frac{4}{5}\right)^N \max_{\frac{1}{2}\bar{D}} |G| \leq \left(\frac{4}{5}\right)^N \sup_D |G| = \left(\frac{4}{5}\right)^N \sup_D |F|,$$

whence, the estimate. \square

4.6 Completing the proof of Theorem 1

Take an arbitrary L -Lipschitz curve $\Gamma = \{t + is(t) : 0 \leq t < \infty\}$, $|s(t_1) - s(t_2)| \leq L|t_1 - t_2|$, and let $\Gamma_K = \{t + is(t) : t \in K\}$ be the part of Γ that lies over an interval K . Since the curve Γ is L -Lipschitz and $c_J + is(c_J) \in \Gamma_J$, we see that Γ_J does not exit the rectangle

$$\{t + iv : t \in J, |v - s(c_J)| \leq \tfrac{1}{2}L|J|\}$$

Let s'_J be a point in \mathfrak{S} closest to $s(c_J)$ (if there are two such points, choose any of them). Put $D_J = D_{J,s'_J}$. Then, $\Gamma_J \subset \bar{D}_J$. Therefore,

$$\Gamma_{\mathbb{J}} \setminus \Pi_{\mathbb{J}_0} = \bigcup_{J \subset \mathbb{J}} \Gamma_J \subset \bigcup_{J \subset \mathbb{J}} \bar{D}_J.$$

By Claim 9, Q does not vanish in the vertical strips $\Pi_{\mathbb{J}_0}$ generated by the untested parts \mathbb{J}_0 . Therefore,

$$N(\Gamma; Q) \leq \sum_{\mathbb{J}} \sum_{J \subset \mathbb{J}} N(\bar{D}_J; Q),$$

where $N(\Gamma; Q)$ is the number of zeroes of Q on Γ .

By Claim 10, $N(\bar{D}_J; Q) \leq C(A, L) \log^2 n$.

At last, recall that the number of intervals J used per each interval \mathbb{J} is at most $CL \log n$, and that the number of the intervals \mathbb{J} where the central index ν stays fixed is at most $V + 1 \leq 2V$. All together, this gives us

$$N(\Gamma; Q) \leq C(A, L)V \log^3 n,$$

completing the proof of Theorem 1. □

References

- [1] A. BLOCH, G. PÓLYA, On the roots of certain algebraic equations. Proc. London Math. Soc. **33** (1932), 102–114.
- [2] P. ERDŐS, A. C. OFFORD, On the number of real roots of a random algebraic equation. Proc. London Math. Soc. (3) **6** (1956), 139–160.
- [3] I. IBRAGIMOV, DM. ZAPOROZHETS, On distribution of zeros of random polynomials in complex plane. In: Prokhorov and contemporary probability theory, 303–323, Springer Proc. Math. Stat., **33**, Springer, Heidelberg, 2013.
- [4] I. A. IBRAGIMOV, N. B. MASLOVA, The mean number of real zeros of random polynomials. I. Coefficients with zero mean (Russian). Teor. Veroyatnost. i Primenen. **16** (1971) 229–248; English transl. in Theor. Probability Appl. **16** (1971), 228–248;
II. Coefficients with a nonzero mean, *ibid*, 495–503; English transl. 485–493.
- [5] I. A. IBRAGIMOV, N. B. MASLOVA, The average number of real roots of random polynomials. (Russian) Dokl. Akad. Nauk SSSR **199** (1971), 13–16; English transl. in Soviet Math. Dokl. **12** (1971), 1004–1008.
- [6] M. KAC, On the average number of real roots of a random algebraic equation. Bull. Amer. Math. Soc. **49** (1943), 314–320; A correction, *ibid*, 938.
- [7] A. N. KOLMOGOROV, Sur les propriétés des fonctions de concentrations de M. P. Lévy. Ann. Inst. H. Poincaré **16** (1958), 27–34.

- [8] J. E. LITTLEWOOD, A. C. OFFORD, On the number of real roots of a random algebraic equation.
I. Journal London Math. Soc. **13** (1938), 288–295;
II. Proc. Cambridge Phil Soc. **35** (1939), 133–148;
III. Rec. Math. [Mat. Sbornik] N.S. **12(54)** (1943). 277–286.
- [9] B. F. LOGAN, L. A. SHEPP, Real zeros of random polynomials.
I, Proc. London Math. Soc. (3) **18** (1968), 29–35;
II, *ibid*, 308–314.
- [10] H. L. MONTGOMERY, Ten Lectures on the Interface Between Analytic Number Theory and Harmonic Analysis. Amer. Math. Soc., 1994.
- [11] F. NAZAROV, Local estimates for exponential polynomials and their applications to inequalities of the uncertainty principle type. (Russian) Algebra & Analiz **5** (1993), no. 4, 3–66; English translation in St. Petersburg Math. J. **5** (1994), 663–717.
- [12] G. PÓLYA, G. SZEGŐ, Problems and Theorems in Analysis II. Reprint of the 1976 English translation. Classics in Mathematics. Springer-Verlag, Berlin, 1998.
- [13] K. SÖZE, Real zeroes of random polynomials, II. Descartes’ rule of signs and anti-concentration on the symmetric group.
- [14] L. SHEPP, K. FARAHMAND, Expected number of real zeros of a random polynomial with independent identically distributed symmetric long-tailed coefficients. Teor. Veroyatn. Primen. **55** (2010), 196–204; Theory Probab. Appl. **55** (2011), 173–181.
- [15] D. N. ZAPOROZHETS, An example of a random polynomial with unusual behavior of the roots. (Russian) Teor. Veroyatn. Primen. **50** (2005), 549–555; English translation in Theory Probab. Appl. **50** (2006), 529–535.

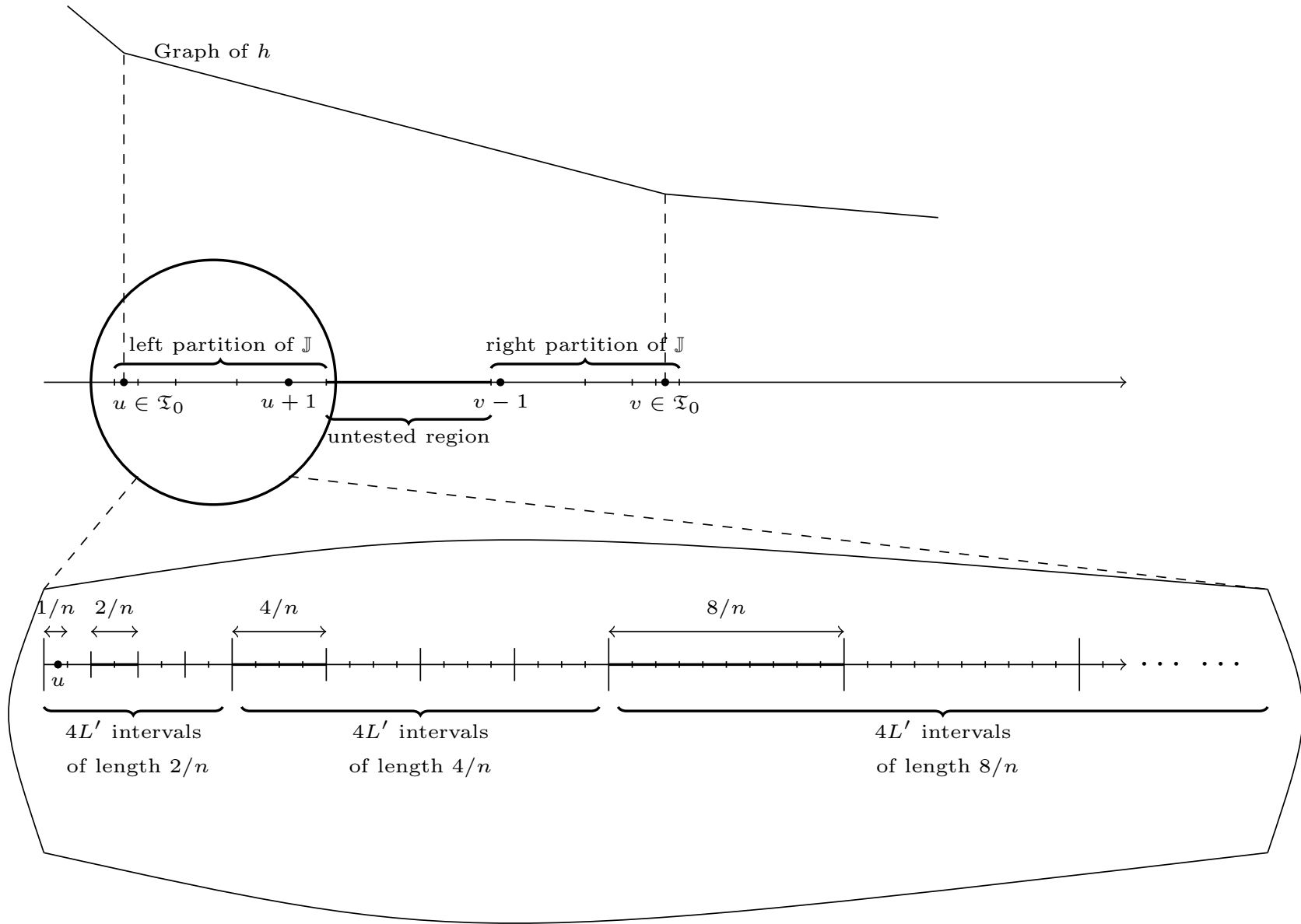


Fig. 1: Partition of the interval $\mathbb{J} = [u, v]$ with $u, v \in \mathfrak{T}_0$; test points are within distance $1/n$ of each other. This figure corresponds to the (impossible) value $L' = 1$.